

15th International Roundtable on Business Survey Frames
Washington, D.C. – October 22 – 26, 2001
<div>Session No 3</div> <div>Paper No 3</div> <div>Koki TOIDA, Statistics Bureau¹, Japan</div>
New Database for the Japanese Business Register

1. Introduction

The Statistics Bureau and the Statistics Center have been maintaining the Establishment and Enterprise Information Database System since January 1998. The database was created for the purpose of providing the lists of establishments and enterprises that are used as the sample frames of the related surveys conducted by the ministries and agencies. It followed the report called "New Strategies for Government Statistical Services for the Coming Decade" issued by the Statistics Council in March 1995.

In April 1999, in order to reduce response burden, the Cabinet decided that the Ministry of Public Management, Home Affairs, Posts and Telecommunications (the Statistics Bureau and the Statistics Center, in practical) should maintain a comprehensive database that is updated by the results of censuses and surveys of the ministries and agencies, and by available administrative sources, and then the ministries and agencies should do the avoidance of the situation that the same object (establishment or enterprise) is selected as samples of a lot of surveys by using the database, before carrying out surveys.

However, the current database has stored the information on only the results of the Establishment and Enterprise Census, and no historical information whether each establishment or enterprise has been selected as the object of the various surveys, which is needed for avoiding the repetition.

So, the Statistics Bureau decided to develop another new database with the present one in cooperation with the Statistics Center in order to meet the request of the Cabinet. The Bureau established the project team composed of six staff members for the system design in December 1999.

The software of the system is under the development by the private company (IBM Japan), and will be completed in October 2001. The hardware was carried in the Bureau at the end of this August. After going through the internal operations, the new database will open to the ministries and agencies for the practical use in April 2002.

2. Roles of the new database

The roles of the new database can be summarised to the following three points.

- a) Preparing and providing the historical information on various censuses and surveys as a tool for avoiding the repetition of survey objects, based on the Cabinet decision
- b) Providing the latest information on the population of establishments and enterprises to be used as the sample frames of the related surveys
- c) Preparing the functions and the data which realize the various analyses on enterprises

2.1 The first point of the roles

In the first point, various bodies have been requesting for avoiding the situation that the same object (establishment or enterprise) is selected as samples of a lot of surveys. With the Cabinet decision as

¹ The views expressed in this paper are those of the author and do not necessarily reflect the policies of the Statistics Bureau of Japan.

a momentum, and from the viewpoint that the surveys should be conducted smoothly, it has come to need to develop a new database.

On the other hand, how to design the framework of avoiding the repetition is a big problem. As for this point, it can be summarized as:

- a) Utilizing the new database for the purpose of conducting the avoidance of the repetition of survey objects efficiently;
- b) Recording the historical information on all the censuses and surveys for the private establishments and enterprises which are conducted by the ministries and agencies into the new database;
- c) Realizing the avoidance of the repetition by the method of selecting a substitutive establishment or enterprise, when the candidate has already met the upper limit of the surveyed number of times.

2.2 The second point of the roles

The Statistics Council has, in the report called "New Strategies for Government Statistical Services for the Coming Decade" in March 1995, indicated that it was necessary to offer the latest information on the population of establishments and enterprises to be used as the sample frames of the related surveys, by using the results of several surveys as the updating information to capture the frequent changes in establishments and enterprises such as newly-established and giving up. Actually, various changes in establishments and enterprises such as the appearance of the establishments having new business forms (e.g. SOHO) caused by the development of IT, and the activation of enterprise reform due to the change of economic structure, have been seen in recent years.

It is the second mission given to the new database to provide the latest information by coping with these changes with the use of various sources for data updating.

2.3 The third point of the roles

The building of the new database itself is the process of avoiding the repetition of survey objects, and is the practical method to realize it. Although the database will record information not on all the results of the censuses but on the necessary minimum for managing the historical information and utilizing as a sampling frame, the information needed for grasping such changes in establishments and enterprises is to be recorded into the database so much as possible.

Activity forms, production systems and decision-making in enterprises have been changing greatly. How the database meets the request for providing the information on these changes in enterprise activities is the third point.

In terms of this, the new database will manage all the establishments and enterprises with the individual codes, which are the establishment code and the enterprise code respectively, and will provide the effective information for the analyses of the structural relations of enterprises and establishments and of the industries of enterprises, etc. Thus, it will be possible to link among surveys by the unification of the establishment code and to use the stored information as the panel information on enterprises.

Note: An 'enterprise' in this paper means a company (single or head office)

3. Characteristics of the new database

The new database that is being developed, based on the above roles and needs, will have necessary characteristics and functions.

3.1 The database of common use type and providing information on-line

This is to develop and manage the database in order to be used commonly by the ministries and agencies for the new subject of avoiding the repetition of survey objects.

From the point of the Cabinet decision, the new database will be a tool for avoiding the situation that the same object (establishment or enterprise) is selected as samples of a lot of surveys, as already

stated. And, the avoidance of the repetition itself is supposed to be done by the ministries and agencies. Therefore, it is important that each ministry or agency can do the exclusion work efficiently. In other words, it is indispensable that the database can be used commonly among the ministries and agencies and makes their efficient operations for avoiding the repetition possible.

Regarding this, it is being cleared that the database will be a common use type among the ministries and agencies and will provide information on-line via the Kasumigaseki-WAN which is the communication network among the ministries and agencies.

3.2 Creation of the establishment code commonly used among the censuses and surveys

The second characteristic of the new database is the creation of the identification code for each establishment as it is commonly used among the related censuses and surveys.

The purposes of creating the establishment code are:

- a) To manage the establishments efficiently and consistently grasped by the various censuses and surveys which have different universes and definitions among them;
- b) To increase the efficiency of the data matching between the database data and the result data of the censuses and surveys needed to update the database and to store the historical information;
- c) To make possible an advanced use of the statistical data on establishments and enterprises.

As for surviving (continuing) establishments especially, the data matching can be done by the establishment code. It will realize not only the efficiency of the data matching and updating but the possible analyses by matching the same establishment among the different surveys.

3.3 Providing the latest information needed to the conduct of surveys

Another character of the new database is to provide the latest information on establishments and enterprises for the ministries and agencies due to the update of the database by using the results of the Census of Commerce, the Census of Manufactures and the Financial Statements of Corporations, in addition to the Establishment and Enterprise Census.

Basic information on establishments and enterprises is prepared by the above censuses separately, and it is difficult to correlate the information mutually. In the new database, however, the items having common definitions among the source data are stored to the common table, while those having the differences are stored separately. By this treatment, it will be possible to refer the information mutually and to prepare the database data having the latest information.

It is of course that the population data recorded in the database will be used as the data for the avoidance of the repetition.

3.4 Use of administrative records and information by private firms

One more thing to take up as the characteristic of the new database is a practical use of the information prepared by private companies.

Regarding the information on enterprises, it is needed to grasp the process from birth to death of an enterprise as a series and to provide it. In other words, it is important to prepare the information on newly-established, merger, takeover, dividing, move and name change of an enterprise, and on the whole figure of enterprise such as the conditions of the parent company and its children and the related companies.

There is a limit, however, in using only the results of the censuses and surveys conducted by the national organizations and in preparing the information on enterprises. Although it must be effective to use the administrative records such as a commercial register and a report on securities, it has not reached to be possible to use them electronically yet. With the consideration of such current status, it has been decided to use the population data of enterprises prepared by the Teikoku Databank (company) as the complimenting source for updating the database.

4. Outline of the new database

4.1 Policies for the development

The following things as the fundamental design policies for the development of the new database have been made, based on the requests for the database indicated above, and with the consideration of proceeding the countermeasure to information security as a part of IT policy by the government.

- a) Securing the system performance
- b) Maintaining the information security
- c) Recovering the system troubles rapidly and properly

Securing the database performance is to keep enough speed and property in on-line data providing via Kasumigaseki-WAN while the database is processed frequently by enormous volume of data. And it is to take a prudent measure against unjust access to the database because it records the individual information on establishments and enterprises.

4.2 System composition of the database

The new database is developed as a client/server system. Connecting the database server and PCs of the ministries and agencies as the clients, it will be possible for the clients to do the on-line-based work for avoiding the repetition.

The system composition of the new database is shown in the ANNEX 1.

The new database system is composed of the population information table group, the historical information table group and the control information table group, and each table group has necessary tables. The information of each table is linked by the establishment code so that the information among the different tables can be connected mutually.

The ANNEX 2 states the concept of the new database.

4.3 Contents of the database

4.3.1 Population data

The population data of the database are prepared by the results of the Establishment and Enterprise Census conducted by the Statistics Bureau, the Census of Commerce and the Census of Manufactures by the Ministry of Economy, Trade and Industry, and the Financial Statements of Corporations by the Ministry of Finance, as the basic updating sources.

The conduction periods of the censuses used to update the population data of the database

Updating sources	Periods of conductions
Establishment and Enterprise Census	In every 5 years (the simplified census is conducted <u>In the middle year between the ordinary censuses</u>)
Census of Commerce	In every 3 years
Census of Manufactures	In every year
Financial Statements of Corporations	In every year (the Statements is prepared by the results of the Survey on Corporation Enterprises)

Except for these sources, the data prepared by the Teikoku Databank are used for the purpose of complimenting the information on enterprises.

Information being recorded in the database can be divided into three kinds of information on private establishments, on establishments under the management of the national and local governments, and on enterprises (companies), roughly.

About 7.5 million establishments will be recorded in the database, consisting of 7.3 million private establishments and 0.2 million governmental ones. The number of the enterprises (companies) will be approximately 1.7 million.

4.3.2 Historical information

The historical information, including the number of times that each establishment has already been surveyed, is prepared by the results of all the censuses and surveys for private establishments and enterprises that are conducted by the ministries and agencies.

The number of censuses and surveys to be used for storing the historical information will be around 230, which are consisted of 13.0 million establishments in total, while the surveys for avoiding the repetition will be 140 and 3.0 million, respectively.

The items on the population data and the historical information to be recorded in the database appear in the ANNEX 3.

4.4 Updating the database

The database data are updated by matching to the data for updating with the use of the establishment code or the name, telephone number and address of the establishment. When only a part of those keys, except for the establishment code, is matched, the persons in charge do a confirmation examination by the visual observation for the purpose of securing the matching accuracy.

4.4.1 Updating the population data

After the matching process, the population data are renewed by the updating data.

- a) The fundamental items (name, address and telephone number of the establishment) of the database data are replaced by the contents of the data for updating.
- b) The attribute items (the industrial classification code, the number of employees, etc.) are recorded in the table of each updating source.
- c) The item on the change of business activities (birth, continuation, death, etc.) is not replaced but stored in the business demography table.

When there are the updating data, which have not existed in the database, the data are newly recorded in the database with the new establishment codes.

4.4.2 Updating the historical information

The matching process itself is same as that for updating the population data. Regarding the historical information, however, the updating means not a renewal but an accumulation. For the matched establishment, in every time when the updating data and the population data of database are matched, the number of times already surveyed in the specific table which is prepared in every establishment and in every month, increases once in the system automatically. The related information on the historical information such as the name of the survey and the name of the ministry which carried out the survey is stored in another table in the database.

When there are the establishments for updating the historical information, which have not existed in the database, the data are recorded in the unmatched data table in the database with the new provisional establishment codes. These establishments will be added to the survey objects of the next Establishment and Enterprise Census.

4.5 Functions of the database

The new database is composed of the following major functions.

4.5.1 Function for data updating the database

- Data matching function

The function to match the population data of the database and the data for updating with the use of the establishment code or the name, telephone number and address of the establishment

- Updating function
The function to update the database data based on the result of matching the population data of the database and the data for updating

4.5.2 Function for using the database

- Searching function
The function to search the needed data from the database
- Tabulating function
The function to tabulate the database data for sampling, the results of avoiding the repetition, etc.
- Sorting function
The function to change data into the order for which users hope
- Downloading function
The function to download the searched data to the client PCs

4.5.3 Function for managing the database

- Function for managing the database data
The function to record and manage the periods of recording, accumulating, and updating the data
- Function for managing the results of use
The function to store and manage the information on the use of the database and on the user's information
- Function for data security
The function to prevent injustice access to the database

4.6 Use of the database

4.6.1 Users and the purpose of use

For the time being, only the ministries and agencies of the national government will be able to use the new database. The database will provide the users with the information on grasping the objects of censuses and surveys, on the avoidance of the repetition, and on analysing the structures and the activities of enterprises.

4.6.2 Avoiding the repetition of survey objects with the use of the database

The avoidance of the situation that the same object (establishment or enterprise) is selected as samples of a lot of surveys is done by the method of selecting an alternative establishment or enterprise, when the candidate for the survey object has already met the upper limit of the surveyed number of times within a year.

The surveys which require for avoiding the repetition are all the statistical surveys for the private establishments or enterprises conducted by the national government, excluding the census. However, all the censuses and surveys targeting the private establishments or enterprises conducted by the national government are used for storing the historical information to the database.

The surveyed number of times is counted as once in every census or survey and in the survey period unit. The legal organizational status, the number of persons engaged, the capital, etc. will be used to create the upper limit of the surveyed number of times within a year for each establishment or enterprise.

The practical process of the avoidance of the repetition with the use of the database is as follows.

- a) Each ministry and agency, in advance of carrying out the survey, presents the list of the candidates for the survey objects to the Statistics Bureau (Statistics Center).
- b) The Statistics Center operates the matching process to the presented list with the database. In this process, the database sets the symbol for avoiding repetition to the establishment or enterprise on the list which has already met the upper limit of the surveyed number of times.
- c) Each ministry and agency selects the alternative for the establishment or enterprise having the symbol, which has been set by the database. In this process, the database provides the information on the candidates for alternation for each ministry and agency on-line.
- d) After the conduct of the survey based on the name list of establishments or enterprises which had completed the avoidance of the repetition, each ministry and agency presents the list to the Statistics Center in order to record the historical information to the database.

The mechanism of avoiding the repetition by using the database is shown in the ANNEX 4.

4.6.3 Method of providing information

The database will provide the information needed for avoiding the repetition on-line in principle, while providing the information on establishments and enterprises for carrying out the censuses and surveys with off-line media of FD, MO, and CMT due to a large volume of data.

4.7 Management of the database

4.7.1 Controlling the users

The users of the database will be managed with the division of general users, special users, persons in charge of data updating, and manager.

4.7.2 Managing the system

The information on the results of using the database, and the data which is necessary for the difficulty restoration countermeasure will be stored for the purpose of the system management.

4.7.3 Security countermeasure

The firewall is employed in order to exclude injustice access to the system. Against the injustice conversion of the database data, the restoration of the data will be made possible by preparing the backup system.

4.8 Start of the new database operation

The new database will start the work for the users in April 2002, following that the population data for building and the historical information on the past one year are being recorded in the database between October 2001 and March 2002.

In this connection, the present database is to be maintained continuously for the time being for the use of various analyses of the results of the Establishment and Enterprise Census in the Statistics Bureau and the Statistics Center.

5. Remained problems

In building the new database, there are some problems that need time to a certain extent for resolving.

5.1 Difference in the definitions among the sources

When the information is managed with centralization in the database by using more than one census result as the database sources, the difference in definitions among censuses becomes a problem. These data such as the industrial classification having the different definitions among the censuses used for updating the new database are to be stored in each census table in the database until they are unified into the same definition in future.

5.2 Preparing the information on enterprises

Recently, a change in enterprises is remarkable in such cases as the activation of enterprise reform reflected by the industrial and economical changes. The ministries and agencies need such information showing these actual conditions of enterprises in order to plan their policies. It is necessary to prepare the related information in the database and provide it to meet the requests by the ministries and agencies.

The forthcoming Establishment and Enterprise Census that is to be conducted November 2001 is supposed to have some new items concerning enterprises. These will be the items to link a parent company and its subsidiary companies such as the name and telephone number of the parent company and the address of the parent company, and the items on the history of company (the change conditions of the company form after the last ordinary census) such as the existences of newly-established, merger, takeover, dividing, move, and the name change.

Although these items will not be perfect to show the actual and change conditions of enterprises, the new database is planning to record these data based on the results of the 2001 Census. The related information prepared by a private company may also be used in the near future.

5.3 Use of administrative records

It is recognized that the use of the administrative records such as a commercial register and a report on securities is effective for expanding the information on enterprises in the database. However, it has not reached to be possible to use them electronically yet.

When these records become electronic media, planned within a few years, the practical study forward the use of the records to the database will be commenced.

6. Future plans

The Statistics Bureau, in cooperation with the Statistics Center, started developing the new database, directly based on the Cabinet decision of building a comprehensive database to provide the information to be used for avoiding the situation that the same object is selected as samples of a lot of surveys by using the database, before carrying out surveys.

However, the new database will not be confined to a tool for avoiding the repetition of survey objects. The goal of the database is to realize the Japanese business register. In order to reach the goal, the strategy consisting of the expansion of database data by using other useful information to update the database including the administrative records, and of building various new functions needed, will be advanced from the next year.

The major practical aims to be realized by the new database and the necessary functions to be prepared in the database are as follows.

- a) Increasing the efficiency of the planning work for statistical survey, by the reduction of response burden as a result of avoiding the repetition of the survey objects and the survey items, by providing the service of sampling operation with the use of the latest data in the database, etc.
- b) Smoothing the execution work of statistical survey over, by providing the labels on which the names and addresses of establishments have been printed for mail survey, by providing the questionnaires on which some answers of the last survey have been printed, etc.
- c) Improving statistical precision, by providing the latest population information on establishments and enterprises, by providing the detailed information on the actual status of establishments and enterprises based on the service of custom-made tabulation, etc.
- d) Preparing statistics (data), by providing the information which is effective for the panel analysis of establishments and enterprises and the analysis on the structure of enterprise group, etc.

The details of the expansion of the new database towards the goal are shown in the ANNEX 5.

7. Conclusion

In April 1999, in order to reduce response burden, the Cabinet decided that the Ministry of Public Management, Home Affairs, Posts and Telecommunications (the Statistics Bureau and the Statistics Center, in practical) should maintain a comprehensive database that is updated by the results of censuses and surveys of the ministries and agencies, and by available administrative sources, and then the ministries and agencies should do the avoidance of the situation that the same object (establishment or enterprise) is selected as samples of a lot of surveys by using the database, before carrying out surveys.

Based on the Cabinet decision, the Statistics Bureau, in cooperation with the Statistics Center, started developing a new database in May 2001, after the system designing by the project team in the Bureau and the Center.

However, the new database will not be confined to a tool for avoiding the repetition of survey objects. The Bureau has a vision of what this database should be. The vision is that the database will realize the Japanese business register of all establishments and enterprises. Practically, it is to make the structure that the ministries and agencies can do all their statistical work on establishments and enterprises only by using this database.

Towards the realization of the vision, the strategy consisting of the expansion of database data by using other useful information to update the database including administrative records and of building various new functions will be advanced from the next year.

The ANNEX 6 shows the future vision of the new database as the Japanese business register.

(ANNEXES)

ANNEX 1 System Composition of the New Database

ANNEX 2 Concept of the New Database

ANNEX 3 Recording Information by the Table in the New Database

ANNEX 4 Mechanism of the Process of Avoiding the Repetition by Using the New Database

ANNEX 5 Expansion of the New Database for the Japanese Business Register

ANNEX 6 Future Vision of the New Database as the Japanese Business Register